PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification [6] : C12N 15/29, 15/82, 15/11, 5/10, C07K 14/415 — A1

(11) International Publication Number: **WO 97/47745**

(43) International Publication Date: 18 December 1997 (18.12.97)

(21) International Application Number: PCT/EP97/03070

(22) International Filing Date: 12 June 1997 (12.06.97)

(30) Priority Data:
PCT/ES96/00130 13 June 1996 (13.06.96) WO
(34) Countries for which the regional or international application was filed: ES et al.

(71) Applicant *(for all designated States except US)*: CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS [ES/ES]; Serrano, 117, E-28006 Madrid (ES).

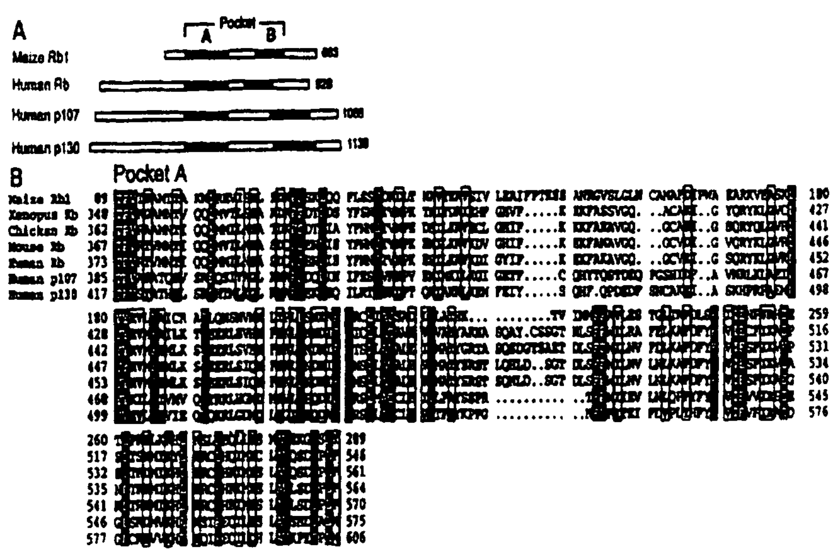
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: GUTIERREZ-ARMENTA, Crisanto [ES/ES]; Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM), Universidad Autónoma, E-28049 Cantoblanco (ES). XIE, Qi [CN/ES]; Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM), Universidad Autónoma, E-28049 Cantoblanco (ES). PELAYO SANZ-BURGOS, Andrés [ES/ES]; Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM), Universidad Autónoma, E-28049 Cantoblanco (ES). SUAREZ LOPEZ, Paula [ES/ES]; Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM), Universidad Autónoma, E-28049 Cantoblanco (ES).

(74) Agent: POLO FLORES, Carlos; Profesor Waksman, 10, E-28036 Madrid (ES).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PLANT RETINOBLASTOMA-ASSOCIATED PROTEINS



(57) Abstract

The present invention is based on the isolation and characterization of a plant cell DNA sequence encoding for a retinoblastoma protein. Such finding is based on the structural and functional properties of the plant retinoblastoma protein as possible regulator of the cellular cycle, of the cellular growth and of the plant cellular differentiation. For this reason, among other aspects, it is claimed the use of retinoblastoma protein or the DNA sequence which encodes for it in the growing control of vegetable cells, plants and/or vegetable virus, as well as the use of vectors, cells, plants or animals, or animal cells modified through the manipulation of the control route based on plant retinoblastoma protein.

Gordon-Kamm *et al.* REF
Serial No. 09/511,445 A2

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

PLANT RETINOBLASTOMA-ASSOCIATED PROTEINS

DESCRIPTION

The present invention relates the proteins having biological activity in plant and animal systems, to polynucleotides encoding for the expression of such proteins, to oligonucleotides for use in identifying and synthesizing these proteins and polynucleotides, to vectors and cells containing the polynucleotides in recombinant form and to plants and animals comprising these, and to the use of the proteins and polynucleotides and fragments thereof in the control of plant growth and plant vulnerability to viruses.

Cell cycle progression is regulated by positive and negative effectors. Among the latter, the product of the retinoblastoma susceptibility gene (Rb) controls the passage of mammalian cells through G1 phase. In mammalian cells, Rb regulates G1/S transit by inhibiting the function of the E2F family of transcription factors, known to interact with sequences in the promoter region of genes required for cellular DNA replication (see eg Weinberg, R.A. Cell 81,323 (1995); Nevins, J.R. Science 258,424 (1992)). DNA tumor viruses that infect animal cells express oncoproteins that interact with the Rb protein via a LXCXE motif, disrupting Rb-E2F complexes and driving cells into S-phase (Weinberg ibid; Ludlow, J. W. FASEB J. 7, 866 (1993); Moran, E. FASEB J. 7, 880 (1993); Vousden, K. FASEB J. 7, 872 (1993)).

The present inventors have shown that efficient replication of a plant geminivirus requires the integrity of an LXCXE amino acid motif in the viral RepA protein and that RepA can interact with members of the human Rb family in yeast (Xie, Q., Suárez-López, P. and Gutiérrez, C. EMBO J. 14, 4073 (1995). The presence of the LXCXE motif in plant D-type cyclins has also been reported (Soni, R., Carmichael, J. P., Shah, Z. H. and Murray, J.

A. H. Plant Cell 7, 85-103 (1995)).

The present inventors have now identified characteristic sequences of plant Rb proteins and corresponding encoding polynucleotides for the first time, isolated such a protein and polynucleotide, and particularly have identified sequences that distinguish it from known animal Rb protein sequences. The inventors have determined that a known DNA sequence from the maize encoding a vegetable Rb plant protein and is hereinafter called ZmRb1. ZmRb1 has been demonstrated by the inventors to interact in yeasts with RepA, a plant geminivirus protein containing LXCXE motif essential for its function. The inventors have further determined that geminivirus DNA replication is reduced in plant cells transfected with plasmids encoding either ZmRb1 or human p130, a member of the human Rb family.

Significantly the inventors work suggests that plant and animal cells may share fundamentally similar strategies for growth control, and thus human as well as plant Rb protein such as ZmRb1 will be expected to have utility in, *inter alia*, plant therapeutics, diagnostics, growth control or investigations and many such plant proteins will have similar utility in animals.

In a first aspect of the present invention there is provided the use of retinoblastoma protein in controlling the growth of plant cells and/or plant viruses. Particularly, the present invention provides control of viral infection and/or growth in plant cells wherein the virus requires the integrity of an LXCXE amino acid motif in one of its proteins, particularly, e. g., in the viral RepA protein, for normal reproduction. Particular plant viruses so controlled are Geminiviruses.

A preferred method of control using such proteins involves applying these to the plant cell, either directly or by introduction of DNA or RNA encoding for

their expression into the plant cell which it is desired to treat. By over expressing the retinoblastoma protein, or expressing an Rb protein or peptide fragment thereof that interacts with the LXCXE motif of the virus but does not affect the normal functioning of the cell, it is possible to inhibit normal virus growth and thus also to produce infection spreading from that cell to its neighbours.

Alternatively, by means of introducing anti-sense DNA or RNA in plant cells in vectors form that contain the necessary promoters for the DNA or RNA transcription, it will be possible to exploit the well known anti-sense mechanism in order to inhibit the expression of the Rb protein, and thus the S-phase. Such plants will be of use, among other aspects to replicate DNA or RNA until high levels, e.g. in yeasts. The methods to introduce anti-sense DNA in cells are very well known for those skilled in the art: see for example "Principles of gene manipulation - An introduction to Genetic Engineering (1994) R.W. Old & S.B. Primrose; Oxford-Blackwell Scientific Publications Fifth Edition p398.

In a second aspect of the present invention there is provided recombinant nucleic acid, particularly in the form of DNA or cRNA (mRNA), encoding for expression of Rb protein that is characteristic of plants. This nucleic acid is characterised by one or more characteristic regions that differ from known animal Rb protein nucleic acid and is exemplified herein by SEQ ID No 1, bases 31-2079.

The DNA or RNA can have a sequence that contains the degenerated substitution in the nucleotides of the codons in SEQ ID No. 1, and in where the RNA the T is U. The most preferred DNA or RNA are capable of hybridate with the polynucleotide of the SEQ ID No. 1 in conditions of low stringency, preferably being the hybridization

produced in conditions of high stringency.

The expressions "conditions of low stringency" and "conditions of high stringency" are understood by those skilled, but are conveniently exemplified in US 5202257, Col-9-Col 10. If some modifications were made to lead to the expression of a protein with different amino acids, preferably of the same kind of the corresponding amino acids to the SEQ ID No 1; that is, are conservative substitutions. Such substitutions are known by those skilled, for example, see US 5380712, and it is only contemplated when the protein has activity with retinoblastoma protein.

Preferred DNA or cRNA encodes for a plant Rb protein having A and B pocket sub-domains having between 30% and 75% homology with human Rb protein, particularly as compared with p130, more preferably from 50% to 64% homology. Particularly the plant Rb protein so encoded has the C706 amino acid of human Rb conserved. Preferably the spacer sequence between the A and B pockets is not conserved with respect to animal Rb proteins, preferably being less than 50% homologous to the same region as found in such animal proteins. Most preferably the protein so encoded has 80% or more homology with that of SEQ NO 2 of the sequence listing attached hereto, still more preferably 90% or more and most preferably 95% or more. Particularly provided is recombinant DNA of SEQ ID No 1 bases 31 to 2079, or the entire SEQ ID No 1, or corresponding RNAs, encoding for maize cDNA clone encoding ZmRb1 of SQ ID No 2.

In a third aspect of the present invention there is provided the protein expressed by the recombinant DNA or RNA of the second aspect, novel proteins derived from such DNA or RNA, and protein derived from naturally occurring DNA or RNA by mutagenic means such as use of mutagenic PCR primers.

In a fourth aspect there are provided vectors, cells and plants and animals comprising the recombinant DNA or RNA of correct sense or anti-sense, of the invention.

In a particularly preferred use of the first aspect there is provided a method of controlling cell or viral growth comprising administering the DNA, RNA or protein of the second or third aspects to the cell. Such administration may be direct in the case of proteins or may involve indirect means, such as electroporation of plant seed cells with DNA or by transformation of cells with expression vectors capable of expressing or over expressing the proteins of the invention or fragments thereof that are capable of inhibiting cell or viral growth.

Alternatively, the method uses an expression vector capable of producing anti-sense RNA of the cDNA of the invention.

Another one of the specific characteristics of the plants protein and of the nucleic acids includes a N-terminal domain corresponding in sequence to the amino acids 1 to 90 of the SEQ ID No. 2 and a nucleotides sequence corresponding to the basis 31 to 300 of the SEQ ID No. 1. These sequences are characterized by possessing less than 150 and less than 450 units that the animal sequences which possess more than 300 amino acids and 900 pairs of more bases.

The present invention will now be illustrated further by reference to the following non-limiting Examples. Further embodiments falling within the scope of the claims attached hereto will occur to those skilled in the light of these.

Figures.

Fig. 1. The sub-figure A shows the relative lengths of the present ZmRb1 protein and the human retinoblastoma proteins. The sub-figure B shows the alignment of the

amino acids sequences of the Pocket A and Pocket B of the ZmRb1 with that of the Xenopus, chicken, rat and three human protein (Rb, p107 and p130).

Fig. 2. This figure is a map of the main characteristics of the WDV virus and the pWori vector derived from WDV and the positions of the deletions and mutations used in order to establish that the LXCXE motif is required for its replication in plants cells.

EXAMPLE 1.

Isolation of DNA and protein expressing clones.

Total RNA was isolated from maize root and mature leaves by grinding the material previously frozen in liquid nitrogen essentially as described in Soni et al (1995). The major and minor p75ZmRb1 mRNAs were identified by hybridization to a random-primed 32P-labelled PstI internal fragment (1.4 kb).

A portion of a maize cDNA library (106 pfu) in lZAPII (Stratagene) was screened by subsequent hybridization to 5'-labelled oligonucleotides designed to be complementary to a known EST sequence of homologue maize of p130. These oligonucleotides were 5'-AATAGACACATCGATCAA/G (M.5m, nt positions 1411-1438) and 5'-GTAATGATACCAACATGG (M.3c, nt positions 1606-1590)(Isogen Biosciences).

After the second round of screening, pBluescript SK-(pBS) phagemids from positive clones were isolated by in vivo excision with ExAssist helper phage (Stratagene) according to protocols recommended by the manufacturer. DNA sequencing was carried out using a SequenaseTM Kit (USB).

The 5'-end of the mRNAs encoding p75ZmRb1 was determined by RACE-PCR. Poly-A+mRNA was purified by chromatography on oligo-dT-cellulose (Amersham). The first strand was synthesized using oligonucleotide DraI35 (5'-GATTTAAAATCAAGCTCC, nt positions 113-96). After denaturation at 90°C for 3 min, RNA was eliminated by

RNase treatment, the cDNA recovered and 5'-tailed with terminal transferase and dATP. Then a PCR fragment was amplified using primer DraI35 and the linker-primer (50 bp) of the Stratagene cDNA synthesis kit.

One of the positive clones so produced contained a ~4 kb insert that, according to restriction analysis, extended both 5' and 3' of the region contained in the Expressed Sequence Tag used. The nucleotide sequence corresponding to the longest cDNA insert (3747 bp) is shown in SEQ ID No. 1. This ZmRb1 cDNA contains a single open reading frame capable of encoding a protein of 683 amino acids (predicted Mr 75247, p75ZmRb1) followed by a 1646 bp 3'-untranslated region. Untranslated regions of similar length have been also found in mammalian Rb cDNAs (Lee, W.-L. et al, Science 235, 1394 (1987); Bernards, R. et al, Proc. Natl. Acad. Sci. USA 86, 6474 (1989)). Northern analysis indicates that maize cells derived from both root meristems and mature leaves contain a major message, ~2.7±0.2 kb in length. In addition, a minor ~3.7±0.2 kb message also appears. Heterogeneous transcripts have been detected in other species (Destrée, O. H. J. et al, Dev. Biol. 153, 141 (1992)).

Plasmid pWoriΔΔ was constructed by deleting in pWori most of the sequences encoding WDV proteins (Sanz and Gutierrez, unpublished). Plasmid p35S.Rb1 was constructed by inserting the CaMV 35S promoter (obtained from pWDV3:35SGUS) upstream of the ZmRb1 cDNA in the pBS vector. Plasmid p35S.130 was constructed by introducing the complete coding sequence of human p130 instead of ZmRb1 sequences into p35S.Rb1. Plasmid p35.A+B was constructed by substituting sequences encoding the WDV RepA and RepB ORFs instead of ZmRb1 in p35S.Rb1 plasmid. (See Soni, R. and Murray, J. A. H. Anal. Biochem. 218, 474-476 (1994)).

The sequence around the methionine codon at nucleotide

position 31 contains a consensus translation start (Kozak, M. J. Mol. Biol. 196, 947 (1987)). To determine whether the cDNA contained the full-length ZmRb1 coding region, the 5'-end of the mRNAs was amplified by RACE-PCR using an oligonucleotide derived from a region close to the putative initiator AUG, which would produce a fragment of ~150 bp. The results are consistent with the ZmRb1 cDNA clone containing the complete coding region.

The ZmRb1 protein contains segments homologous to the A and B subdomains of the "pocket" that is present in all members of the Rb family. These subdomains are separated by a non-conserved spacer. ZmRb1 also contains non-conserved N-terminal and C-terminal domains. Overall, ZmRb1 shares ~28-30% amino acid identity (~50% similarity) with the Rb family members (Hannon, G. J., Demetrick, D. & Beach, D. Genes Dev. 7, 2378 (1993); Cobrinik, D., Whyte, P., Peeper, D.S., Jacks, T. & Weinberg, R. A. ibid., p. 2392 (1993). Ewen, M. E., Xing, Y. Lawrence, J. B. and Livingston, D. M. Cell 66, 1155 (1991))(Lee W. L. et al, Science 235, 1394 (1987); Bernards et al, Proc. Natl. Acad. Sci. USA 86, 6974 (1989)), with the A and B subdomains exhibiting the highest homology (~50-64%). Interestingly, amino acid C706 in human Rb, critical for its function (Kaye, F. J., Kratzke R. A., Gerster, J. L. and Horowitz, J. M. Proc. Natl. Acad. Sci. USA 87, 6922 (1990)), is also conserved in maize p75ZmRb1.

Note: The 561-577 amino acids encompass a proline-rich domain.

ZmRb1 contains 16 consensus sites, SP or TP for phosphorilation by cyclins dependant kinases (CDKs) with one of the 5'-tail of the sub-domain A and several in the C-terminal area which are potential sites of phosphorilation. A nucleic acid preferred group which encodes proteins in which one or more of these sites are

changed or deleted, making the protein more resistant to the phosphorilation and thus, to its functionality, for example linking to E2F or similar. This can be easily carried out by means of mutagenesis conducted by means of PCR.

EXAMPLE 2

In vivo activity.

Replication of wheat dwarf geminivirus (WDV) is dependent upon an intact LXCXE motif of the viral RepA protein. This motif can mediate interaction with a member of the human Rb family, p130, in yeasts. Therefore, the inventors investigated whether p75ZmRb1 could complex with WDV RepA by using the yeast two-hybrid system (Fields, S. and Song, O. Nature 340, 245-246 (1989)). Yeast cells were co-transformed with a plasmid encoding the fusion GAL4BD-RepA protein and with plasmids encoding different GAL4AD fusion protein. The GAL4AD-p75ZmRb1 fusion could also complex with GAL4BD-RepA to allow growth of the recipient yeast cells in the absence of histidine. This interaction was slightly stronger than that seen with the human p130 protein. RepA could also bind to some extent to a N-terminally truncated form of p75ZmRb1. The role of the LXCXE motif in RepA-p75ZmRb1 interaction was assessed using a point mutation in WDV RepA (E198K) which we previously showed to destroy interaction with human p130. Co-transformation of ZmRb1 with a plasmid encoding the fusion GAL4BD-RepA(E198K) indicated that the interaction between RepA and p75ZmRb1 occurred through the LXCXE motif.

In this respect, the E198K mutant of WDV RepA behaves similarly to analogous point mutants of animal virus oncoproteins (Moran, E., Zerler, B., Harrison, T. M. and Mathews, M.B. Mol. Cell Biol. 6, 3470 (1986); Cherington, V. et al., ibid., p. 1380 (1988); Lillie, J. W., Lowenstein, P. M., Green, M. R. and Green, M. Cell 50,

1091 (1987); DeCarpio, J. A. et al., ibid., p. 275 (1988)).

Specific interaction between maize p75ZmRb1 and WDV RepA in the yeast two-hybrid system (Fields et al) relied on the ability to reconstitute a functional GAL4 activity from two separated GAL4 fusion proteins containing the DNA binding domain (GAL4BD) and the activation domain (GAL4AD). Yeast HF7c cells were co-transformed with a plasmid expressing the GAL4BD-RepA or the GAL4BD-RepA(E198K) fusions and the plasmids expressing the GAL4AD alone (Vec) or fused to human p130, maize p75 (p75ZmRb1) or a 69 amino acids N-terminal deletion of p75 (p75ZmRb1-DN). Cells were streaked on plates with or without histidine according to the distribution shown in the upper left corner. The ability to grow in the absence of histidine depends on the functional reconstitution of a GAL4 activity upon interaction of the fusion proteins, since this triggers expression of the HIS3 gene which is under the control of a GAL4 responsive element. The growth characteristics of these yeast co-transformants correlate with the levels of b-galactosidase activity.

Procedures for two-hybrid analysis are described in Xie et al (1995). The GAL4AD-ZmRb1 fusions were construed in the pGAD424 vector.

EXAMPLE 3

In vivo activity.

Geminivirus DNA replication requires the cellular DNA replication machinery as well as other S-phase specific factors (Davies, J. W. and Stanley, J. Trends Genet. 5, 77 (1989); Lazarowitz, S. Crit. Rev. Plant Sci. 11, 327 (1992)). Consistent with this requirement, geminivirus infection appears to drive non-proliferating cells into S-phase, as indicated by the accumulation of the proliferating cell nuclear antigen (PCNA), a protein which is not normally present in the nuclei of

differentiated cells (Nagar, S., Pedersen, T. J., Carrick, K. M., Hanley-Bowdoin, L. and Robertson, D. Plant Cell 7, 705 (1995)). The inventors finding that efficient WDV DNA replication requires an intact LXCXE motif in RepA coupled with the discovery of a plant homolog of Rb supports the model that, as in animal cells, sequestration of plant Rb by viral RepA protein promotes inappropriate entry of infected cells into S-phase. Therefore, one way to investigate the function of p75ZmRb1 was to measure geminivirus DNA replication in cells transfected with a plasmid bearing the ZmRb1 sequences under a promoter functional in plant cells, an approach analogous to that previously used in human cells (Uzvolgi, E. et al., Cell Growth Diff 2, 297 (1991)). Accumulation of newly replicated viral plasmid DNA was impaired in wheat cells transfected with plasmids expressing p75ZmRb1 or human p130, when expression of WDV replication protein(s) is directed wither by the WDV promoter or by the CaMV 35S promoter.

Since WDV DNA replication requires an S-phase cellular environment, interference with viral DNA replication by p75ZmRb1 and human p130 strongly evidences a role for retinoblastoma protein in the control of the G1/S transition in plants. The existence of a plant Rb homolog implies that despite their ancient divergence, plant and animal cells use, at least in part, similar regulatory proteins and pathways for cell cycle control.

Two lines of evidences reinforce this model. First, a gene encoding a protein that complements specifically the G1/S, but not the G2/M transition of the budding yeast cdc28 mutant has been identified in alfalfa cells (Hirt, H., Páy, A., Bögre, L., Meskiene, I. and Heberle-Bors, E. Plant J. 4, 61 (1993)). Second, plant homologs of D-type cyclins have been isolated from Arabidopsis and these, like their mammalian relatives, contain LXCXE motifs. In

concert with plant versions of CDK4 and CDK6, plant D-type cyclins may regulate passage through G1 phase by controlling the phosphorylation state of Rb-like proteins.

In animal cells, the Rb family has been implicated in tumor suppression and in the control of differentiation and development. Thus, p75ZmRb1 could also play key regulatory roles at other levels during the plant cell life. One key question that is raised by the existence of Rb homologs in plant cells in whether, as in animals disruption of the Rb pathway leads to a tumor-prone condition. In this regard, the inventors have noted that the VirB4 protein encoded by the Ti plasmids of both *Agrobacterium tumefaciens* and *A. rhyzogenes* contains an LXCXE motif. Although the VirB4 protein is required for tumor induction (Hooykas, P. J. J. and Beijersbergen, A. G. M. Annu. Rev. Phytopathol. 32, 157 (1994), the function of its LXCXE motif in this context remains to be examined. Geminivirus infection is not accompanied by tumor development in the infected plant, but in some cases an abnormal growth of enactions has been observed (G. Dafalla and B. Gronenborn, personal communication).

Inhibition of wheat dwarf geminivirus (WDV) DNA replication by ZmRb1 or human p130 in cultured wheat cells was carried out as follows. A. Wheat cells were transfected, as indicated, with pWori (Xie et al. 1995) alone (0.5g), a replicating WDV-based plasmid which encodes WDV proteins required for viral DNA replication, and with control plasmid pBS (10 g) or p35S.Rb1 (10 g), which encodes ZmRb1 sequences under the control of the CaMV 35S promoter. Total DNA was purified one and two days after transfection, equal amounts fractionated in agarose gels and ethidium bromide staining and viral pWori DNA identified by Southern hybridization. Plasmid DNA represents exclusively newly-replicated plasmid DNA

since it is fully resistant to DpnI digestion and sensitive to Mbol. Note that the MboI-digested samples were run for about half of the length than the undigested samples. B. To test the effect of human p130 on WDV DNA replication, wheat cells were co-transfected with pWori (0.5 g) and plasmids pBS (control), p35S.Rb1 or p35S.130 (10 g in each case). Replication of the test plasmid (pWori) was analyzed two days after transfection and was detected as described in part A using ethidium bromide staining; and Southern hybridization. C. To test the effect of ZmRb1 or human p130 on WDV DNA replication when expression of viral proteins was directed by the CaMV 35S promoter, the test plasmid pWoriΔΔ (which does not encode functional WDV replication proteins but replicates when they are provided by a different plasmid, i. e. pWori) was used. Wheat cells were co-transfected, as indicated, with pWoriΔΔ (0.25 g), pWori (0.25 g), p35S.A+B (6 g), p35S.Rb1 (10 g) and/or p35S.130 (10 g). Replication of the test plasmid (pWoriΔΔ) was analyzed 36 hours after transfection and was detected as described in part A using ethidium bromide staining; Southern hybridization. Plasmids pWori (M1) and pWoriΔΔ (M2; Sanz and Gutiérrez, unpublished), 100 pg in each case, were used as markers. Suspension cultures of wheat cells, transfection by particle bombardment and analysis of viral DNA replication were carried out as described in (Xie et al. 1995), except that DNA extraction was modified as in (Soni and Murray. Arnal. Biochem. 218, 474-476 (1995).

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: CRISANTO GUTIERREZ ARMENTA

(A) NAME: QI XIE

(A) NAME: ANDRES PELAYO SANZ-BURGOS

(A) NAME: PAULA SUAREZ-LOPEZ

(B) STREET: CSIC-UAM, UNIVERSIDAD AUTONOMA, CANTOBLANCO

(C) CITY: MADRID

(E) COUNTRY: SPAIN

(F) POSTAL CODE (ZIP): 28049

(ii) TITLE OF THE INVENTION: PLANT PROTEINS

(iii) NUMBER OF SEQUENCES: 2

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3747 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Zea mays

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 31..2079

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
GAATTCGGCA CGAGCAAAGG TCTGATTGAT ATG GAA TGT TTC CAG TCA AAT TTG      54
                                 Met Glu Cys Phe Gln Ser Asn Leu
                                  1               5

GAA AAA ATG GAG AAA CTA TGT AAT TCT AAT AGC TGT AAA GGG GAG CTT      102
Glu Lys Met Glu Lys Leu Cys Asn Ser Asn Ser Cys Lys Gly Glu Leu
     10                  15                  20

GAT TTT AAA TCA ATT TTG ATC AAT AAT GAT TAT ATT CCC TAT GAT GAG      150
Asp Phe Lys Ser Ile Leu Ile Asn Asn Asp Tyr Ile Pro Tyr Asp Glu
 25                  30                  35                  40

AAC TCG ACG GGG GAT TCC ACC AAT TTA GGA CAT TCA AAG TGT GCC TTT      198
Asn Ser Thr Gly Asp Ser Thr Asn Leu Gly His Ser Lys Cys Ala Phe
                 45                  50                  55

GAA ACA TTG GCA TCT CCC ACA AAG ACA ATA AAG AAC ATG CTG ACT GTT      246
Glu Thr Leu Ala Ser Pro Thr Lys Thr Ile Lys Asn Met Leu Thr Val
             60                  65                  70

CCT AGT TCT CCT TTG TCA CCA GCC ACC GGT GGT TCA GTC AAG ATT GTG      294
Pro Ser Ser Pro Leu Ser Pro Ala Thr Gly Gly Ser Val Lys Ile Val
         75                  80                  85

CAA ATG ACA CCA GTA ACT TCT GCC ATG ACG ACA GCT AAG TGG CTT CGT      342
Gln Met Thr Pro Val Thr Ser Ala Met Thr Thr Ala Lys Trp Leu Arg
     90                  95                 100

GAG GTG ATA TCT TCA TTG CCA GAT AAG CCT TCA TCT AAG CTT CAG CAG      390
Glu Val Ile Ser Ser Leu Pro Asp Lys Pro Ser Ser Lys Leu Gln Gln
105                 110                 115                 120

TTT CTG TCA TCA TGC GAT AGG GAT TTG ACA AAT GCT GTC ACA GAA AGG      438
Phe Leu Ser Ser Cys Asp Arg Asp Leu Thr Asn Ala Val Thr Glu Arg
                125                 130                 135

GTC AGC ATA GTT TTG GAA GCA ATT TTT CCA ACC AAA TCT TCT GCC AAT      486
Val Ser Ile Val Leu Glu Ala Ile Phe Pro Thr Lys Ser Ser Ala Asn
            140                 145                 150

CGG GGT GTA TCG TTA GGT CTC AAT TGT GCA AAT GCC TTT GAC ATT CCG      534
Arg Gly Val Ser Leu Gly Leu Asn Cys Ala Asn Ala Phe Asp Ile Pro
        155                 160                 165

TGG GCA GAA GCC AGA AAA GTG GAG GCT TCC AAG TTG TAC TAT AGG GTA      582
Trp Ala Glu Ala Arg Lys Val Glu Ala Ser Lys Leu Tyr Tyr Arg Val
    170                 175                 180
```

TTA GAG GCA ATC TGC AGA GCG GAG TTA CAA AAC AGC AAT GTA AAT AAT 630
Leu Glu Ala Ile Cys Arg Ala Glu Leu Gln Asn Ser Asn Val Asn Asn
185 190 195 200

CTA ACT CCA TTG CTG TCA AAT GAG CGT TTC CAC CGA TGT TTG ATT GCA 678
Leu Thr Pro Leu Leu Ser Asn Glu Arg Phe His Arg Cys Leu Ile Ala
205 210 215

TGT TCA GCG GAC TTA GTA TTG GCG ACA CAT AAG ACA GTC ATC ATG ATG 726
Cys Ser Ala Asp Leu Val Leu Ala Thr His Lys Thr Val Ile Met Met
220 225 230

TTT CCT GCT GTT CTT GAG AGT ACC GGT CTA ACT GCA TTT GAT TTG AGC 774
Phe Pro Ala Val Leu Glu Ser Thr Gly Leu Thr Ala Phe Asp Leu Ser
235 240 245

AAA ATA ATT GAG AAC TTT GTG AGA CAT GAA GAG ACC CTC CCA AGA GAA 822
Lys Ile Ile Glu Asn Phe Val Arg His Glu Glu Thr Leu Pro Arg Glu
250 255 260

TTG AAA AGG CAC CTA AAT TCC TTA GAA GAA CAG CTT TTG GAA AGC ATG 870
Leu Lys Arg His Leu Asn Ser Leu Glu Glu Gln Leu Leu Glu Ser Met
265 270 275 280

GCA TGG GAG AAA GGT TCA TCA TTG TAT AAC TCA CTG ATT GTT GCC AGG 918
Ala Trp Glu Lys Gly Ser Ser Leu Tyr Asn Ser Leu Ile Val Ala Arg
285 290 295

CCA TCT GTT GCT TCA GAA ATA AAC CGC CTT GGT CTT TTG GCT GAA CCA 966
Pro Ser Val Ala Ser Glu Ile Asn Arg Leu Gly Leu Leu Ala Glu Pro
300 305 310

ATG CCA TCT CTT GAT GAC TTA GTG TCA AGG CAG AAT GTT CGT ATC GAG 1014
Met Pro Ser Leu Asp Asp Leu Val Ser Arg Gln Asn Val Arg Ile Glu
315 320 325

GGC TTG CCT GCT ACA CCA TCT AAA AAA CGT GCT GCT GGT CCA GAT GAC 1062
Gly Leu Pro Ala Thr Pro Ser Lys Lys Arg Ala Ala Gly Pro Asp Asp
330 335 340

AAC GCT GAT CCT CGA TCA CCA AAG AGA TCG TGC AAT GAA TCT AGG AAC 1110
Asn Ala Asp Pro Arg Ser Pro Lys Arg Ser Cys Asn Glu Ser Arg Asn
345 350 355 360

ACA GTA GTA GAG CGC AAT TTG CAG ACA CCT CCA CCC AAG CAA AGC CAC 1158
Thr Val Val Glu Arg Asn Leu Gln Thr Pro Pro Pro Lys Gln Ser His
365 370 375

ATG GTG TCA ACT AGT TTG AAA GCA AAA TGC CAT CCA CTC CAG TCC ACA 1206
Met Val Ser Thr Ser Leu Lys Ala Lys Cys His Pro Leu Gln Ser Thr
380 385 390

TTT GCA AGT CCA ACT GTC TGT AAT CCT GTT GGT GGG AAT GAA AAA TGT 1254

Phe Ala Ser Pro Thr Val Cys Asn Pro Val Gly Gly Asn Glu Lys Cys
395 400 405

GCT GAC GTG ACA ATT CAT ATA TTC TTT TCC AAG ATT CTG AAG TTG GCT 1302
Ala Asp Val Thr Ile His Ile Phe Phe Ser Lys Ile Leu Lys Leu Ala
410 415 420

GCT ATT AGA ATA AGA AAC TTG TGC GAA AGG GTT CAA TGT GTG GAA CAG 1350
Ala Ile Arg Ile Arg Asn Leu Cys Glu Arg Val Gln Cys Val Glu Gln
425 430 435 440

ACA GAG CGT GTC TAT AAT GTC TTC AAG CAG ATT CTT GAG CAA CAG ACA 1398
Thr Glu Arg Val Tyr Asn Val Phe Lys Gln Ile Leu Glu Gln Gln Thr
445 450 455

ACA TTA TTT TTT AAT AGA CAC ATC GAT CAA CTT ATC CTT TGC TGT CTT 1446
Thr Leu Phe Phe Asn Arg His Ile Asp Gln Leu Ile Leu Cys Cys Leu
460 465 470

TAT GGT GTT GCA AAG GTT TGT CAA TTA GAA CTC ACA TTC AGG GAG ATA 1494
Tyr Gly Val Ala Lys Val Cys Gln Leu Glu Leu Thr Phe Arg Glu Ile
475 480 485

CTC AAC AAT TAC AAA AGA GAA GCA CAA TGC AAG CCA GAA GTT TTT TCA 1542
Leu Asn Asn Tyr Lys Arg Glu Ala Gln Cys Lys Pro Glu Val Phe Ser
490 495 500

AGT ATC TAT ATT GGG AGT ACG AAC CGT AAT GGG GTA TTA GTA TCG CGC 1590
Ser Ile Tyr Ile Gly Ser Thr Asn Arg Asn Gly Val Leu Val Ser Arg
505 510 515 520

CAT GTT GGT ATC ATT ACT TTT TAC AAT GAG GTA TTT GTT CCA GCA GCG 1638
His Val Gly Ile Ile Thr Phe Tyr Asn Glu Val Phe Val Pro Ala Ala
525 530 535

AAG CCT TTC CTG GTG TCA CTA ATA TCA TCT GGT ACT CAT CCA GAA GAC 1686
Lys Pro Phe Leu Val Ser Leu Ile Ser Ser Gly Thr His Pro Glu Asp
540 545 550

AAG AAG AAT GCT AGT GGC CAA ATT CCT GGA TCA CCC AAG CCA TCT CCT 1734
Lys Lys Asn Ala Ser Gly Gln Ile Pro Gly Ser Pro Lys Pro Ser Pro
555 560 565

TTC CCA AAT TTA CCA GAT ATG TCC CCG AAG AAA GTT TCA GCA TCT CAT 1782
Phe Pro Asn Leu Pro Asp Met Ser Pro Lys Lys Val Ser Ala Ser His
570 575 580

AAT GTA TAT GTG TCT CCT TTG CGG CAA ACC AAG TTG GAT CTA CTG CTG 1830
Asn Val Tyr Val Ser Pro Leu Arg Gln Thr Lys Leu Asp Leu Leu Leu
585 590 595 600

TCA CCA AGT TCC AGG AGT TTT TAT GCA TGC ATT GGT GAA GGC ACC CAT 1878
Ser Pro Ser Ser Arg Ser Phe Tyr Ala Cys Ile Gly Glu Gly Thr His
605 610 615

GCT TAT CAG AGC CCA TCT AAG GAT TTG GCT GCT ATA AAT AGC CGC CTA 1926
Ala Tyr Gln Ser Pro Ser Lys Asp Leu Ala Ala Ile Asn Ser Arg Leu
620 625 630

AAT TAT AAT GGC AGG AAA GTA AAC AGT CGA TTA AAT TTC GAC ATG GTG 1974
Asn Tyr Asn Gly Arg Lys Val Asn Ser Arg Leu Asn Phe Asp Met Val
635 640 645

AGT GAC TCA GTG GTA GCC GGC AGT CTG GGC CAG ATA AAT GGT GGT TCT 2022
Ser Asp Ser Val Val Ala Gly Ser Leu Gly Gln Ile Asn Gly Gly Ser
650 655 660

ACT TCG GAT CCT GCA GCT GCA TTT AGC CCC CTT TCA AAG AAG AGA GAG 2070
Thr Ser Asp Pro Ala Ala Ala Phe Ser Pro Leu Ser Lys Lys Arg Glu
665 670 675 680

ACA GAT ACT TGATCAATTA TAAATGGTGG CCTCTCTCGT ATATAGCTCA 2119
Thr Asp Thr

CAGATCCGTG CTCCGTAGCA GTCTATTCTT CTGAATAAGT GGATTAACTG GAGCGATTTA 2179

ACTGTACATG TATGTGTTAG TGAGAAGCAG CAGTTTTTAG GCAGCAAACT GTTTCAAGTT 2239

AGCTTTTGAG CTATCACCAT TTCTCTGCTG ATTGAACATA TCCGCTGTGT AGAGTGCTAA 2299

TGAATCTTTA GTTTTCATTG GGCTGACATA ACAAATCTTT ATCCTAGTTG GCTGGTTGTT 2359

GGGAGGCATT CATCAGGGTT ATATTTGGTT GTCAAAAAGT ACTGTACTTA ATTCACATCT 2419

TTCACATTTT TCACTAGCAA TAGCAGCCCC AAATTGCTTT CCTGACTAGG AACATATTCT 2479

TTACAGGTAT AAGCATGCCA ACTCTAAACT ATATGAATCC TTTTTATATT CTCATTTTTA 2539

AGTACTTCTC TGTTTCTGCT ACTTTTGTAC TGTATATTTC CAGCTTCTCC ATCAGACTGA 2599

TGATCCCATA TTCAGTGTGC TGCAAGTGAT TTGACCATAT GTGGCTTATC CTTCAGGTAT 2659

GTCTCATGTT GTGACTTCAT TGCTGATTGC TTTTGTAATG GTACTGTTGA GTTCATTTCT 2719

GGTTACAATC AGCCTTTACT GCTTTATATT GTTCTACTAA TTTTGGCTTG CACAGCCAGG 2779

ACGATTGGTT TTCTGCATCA ATCAATCTTT TTTAGGACAA GATATTTTTG TATGCTACAC 2839

TTCCCAAATT GCAATTAATC CAGAAGTCTA CCTTGTTTTA TTCTATTAGT TCTCAGCAAC 2899

AGTGAATGAA TATGAATCAG TCATGCTGAT AGATGTTCAT CTGGTTATTC CAAACAATCT 2959

GACATCGCAT CTCTTTCTGC AAGTGAGATG AAGAAAACCT GAAATGCTAT CACCATTTAA 3019

AACATTGGCT TCTGGAAGTT CAGGTGATTA GCAGGAGACG TTCTGACATT GCCATTGACA 3079

TGTACGGTAG TGATGGCAGG AGACGTTCTT AAACAGCAGC TGCTCCTTCA GCTTGTAATG 3139

TCTGATTGTA TTGACCAAGA GCATCCACCT TGCCTTATGG TACTAACTGA ATGAGCTGGT 3199

GACGCTGACT CATCTGCATA ATGGCAGATG CTTAACCATC TTTAGGAGCT CATGTCATGA 3259

TTCCAGCTGC ACCGTGTCAA ATGTGAAGGC CCTGCAAGGC TTTCCAGGCC GCACCAATCC 3319

TGCTTGCTTC TTGAAGATAC ATATGGTGCC ACCTAAATAA AAGCTGTTTC TGGTTATGTC 3379

TGTCCTTGAC ATGTCAACAG ATTAGTGTTG GGTTGCAGTC ATGTGGTGTT TAAGTCTTGG 3439

AGAAGGCGAG AAGTCATTGC TGCCAGCATT GTGATCGTCA GGCACAGAAG TACTCAAAAG 3499

TGAGAGCTAC TTGTTGCGAG CAAACGGAGG GCGATATAGG TTGATAGCCA ATTTCAGTTC 3559

TCTATATACA AGCAGCGGAT TTTGTTTAGA GTTAGCTTTT GAGATGCATC ATTTCTTTCA 3619

CATCTGATTC TGTGTGTTGT AACTCGGAGT CGCGTAGAAG TTAGAATGCT AACTGACCTT 3679

AATTTTCACC GAATAATTTG CTAGCGTTTT TCAGTATGAA ATCCTTGTCT TAAAAAAAAA 3739

AAAAAAAA 3747

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 683 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Glu Cys Phe Gln Ser Asn Leu Glu Lys Met Glu Lys Leu Cys Asn
1 5 10 15

Ser Asn Ser Cys Lys Gly Glu Leu Asp Phe Lys Ser Ile Leu Ile Asn
20 25 30

Asn Asp Tyr Ile Pro Tyr Asp Glu Asn Ser Thr Gly Asp Ser Thr Asn
35 40 45

Leu Gly His Ser Lys Cys Ala Phe Glu Thr Leu Ala Ser Pro Thr Lys
50 55 60

Thr Ile Lys Asn Met Leu Thr Val Pro Ser Ser Pro Leu Ser Pro Ala
65 70 75 80

Thr Gly Gly Ser Val Lys Ile Val Gln Met Thr Pro Val Thr Ser Ala
85 90 95

Met Thr Thr Ala Lys Trp Leu Arg Glu Val Ile Ser Ser Leu Pro Asp
100 105 110

Lys Pro Ser Ser Lys Leu Gln Gln Phe Leu Ser Ser Cys Asp Arg Asp
115 120 125

Leu Thr Asn Ala Val Thr Glu Arg Val Ser Ile Val Leu Glu Ala Ile
130 135 140

Phe Pro Thr Lys Ser Ser Ala Asn Arg Gly Val Ser Leu Gly Leu Asn
145 150 155 160

Cys Ala Asn Ala Phe Asp Ile Pro Trp Ala Glu Ala Arg Lys Val Glu
165 170 175

Ala Ser Lys Leu Tyr Tyr Arg Val Leu Glu Ala Ile Cys Arg Ala Glu
180 185 190

Leu Gln Asn Ser Asn Val Asn Asn Leu Thr Pro Leu Leu Ser Asn Glu
195 200 205

Arg Phe His Arg Cys Leu Ile Ala Cys Ser Ala Asp Leu Val Leu Ala
210 215 220

Thr His Lys Thr Val Ile Met Met Phe Pro Ala Val Leu Glu Ser Thr
225 230 235 240

Gly Leu Thr Ala Phe Asp Leu Ser Lys Ile Ile Glu Asn Phe Val Arg
245 250 255

His Glu Glu Thr Leu Pro Arg Glu Leu Lys Arg His Leu Asn Ser Leu
260 265 270

Glu Glu Gln Leu Leu Glu Ser Met Ala Trp Glu Lys Gly Ser Ser Leu
275 280 285

Tyr Asn Ser Leu Ile Val Ala Arg Pro Ser Val Ala Ser Glu Ile Asn
290 295 300

Arg Leu Gly Leu Leu Ala Glu Pro Met Pro Ser Leu Asp Asp Leu Val
305 310 315 320

Ser Arg Gln Asn Val Arg Ile Glu Gly Leu Pro Ala Thr Pro Ser Lys
325 330 335

Lys Arg Ala Ala Gly Pro Asp Asp Asn Ala Asp Pro Arg Ser Pro Lys
340 345 350

Arg Ser Cys Asn Glu Ser Arg Asn Thr Val Val Glu Arg Asn Leu Gln
355 360 365

Thr Pro Pro Pro Lys Gln Ser His Met Val Ser Thr Ser Leu Lys Ala
370 375 380

Lys Cys His Pro Leu Gln Ser Thr Phe Ala Ser Pro Thr Val Cys Asn
385 390 395 400

Pro Val Gly Gly Asn Glu Lys Cys Ala Asp Val Thr Ile His Ile Phe
405 410 415

Phe Ser Lys Ile Leu Lys Leu Ala Ala Ile Arg Ile Arg Asn Leu Cys
420 425 430

Glu Arg Val Gln Cys Val Glu Gln Thr Glu Arg Val Tyr Asn Val Phe
435 440 445

Lys Gln Ile Leu Glu Gln Glu Thr Thr Leu Phe Phe Asn Arg His Ile
450 455 460

Asp Gln Leu Ile Leu Cys Cys Leu Tyr Gly Val Ala Lys Val Cys Gln
465 470 475 480

Leu Glu Leu Thr Phe Arg Glu Ile Leu Asn Asn Tyr Lys Arg Glu Ala
485 490 495

Gln Cys Lys Pro Glu Val Phe Ser Ser Ile Tyr Ile Gly Ser Thr Asn
500 505 510

Arg Asn Gly Val Leu Val Ser Arg His Val Gly Ile Ile Thr Phe Tyr
515 520 525

Asn Glu Val Phe Val Pro Ala Ala Lys Pro Phe Leu Val Ser Leu Ile
530 535 540

Ser Ser Gly Thr His Pro Glu Asp Lys Lys Asn Ala Ser Gly Gln Ile
545 550 555 560

Pro Gly Ser Pro Lys Pro Ser Pro Phe Pro Asn Leu Pro Asp Met Ser
565 570 575

Pro Lys Lys Val Ser Ala Ser His Asn Val Tyr Val Ser Pro Leu Arg
580 585 590

Gln Thr Lys Leu Asp Leu Leu Leu Ser Pro Ser Ser Arg Ser Phe Tyr
595 600 605

Ala Cys Ile Gly Glu Gly Thr His Ala Tyr Gln Ser Pro Ser Lys Asp
610 615 620

Leu Ala Ala Ile Asn Ser Arg Leu Asn Tyr Asn Gly Arg Lys Val Asn
625 630 635 640

Ser Arg Leu Asn Phe Asp Met Val Ser Asp Ser Val Val Ala Gly Ser
645 650 655

Leu Gly Gln Ile Asn Gly Gly Ser Thr Ser Asp Pro Ala Ala Ala Phe
660 665 670

Ser Pro Leu Ser Lys Lys Arg Glu Thr Asp Thr
675 680

INFORMATION RELATIVE TO THE DEPOSIT OF A MICRO-ORGANISM

The micro-organism to which reference is made in page 6 of the disclosure has been deposited in the following institution:

COLECCION ESPAÑOLA DE CULTIVOS TIPO (CECT)
Departamento de Microbiología
Facultad de Ciencias Biológicas
46100 EURJASOT (Valencia)
Spain

Deposit identification: pBS.Rb1
Deposit date: June 12, 1996
Order No.: 4699

This information appears reflected in the form PCE/RO/134 enclosed to the request.

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 6, line 24 and following.

B. IDENTIFICATION OF DEPOSIT pBS.Rb1

Further deposits are identified on an additional sheet ☐

Name of depositary institution

COLECCION ESPAÑOLA DE CULTIVOS TIPO (CECT)

Address of depositary institution *(including postal code and country)*

Departamento de Microbiología
Facultad de Ciencias Biológicas
46100 BURJASOT (Valencia)
Spain

Date of deposit	Accession Number
12 June 1996	4699

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)*

This information is continued on an additional sheet ☐

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☒ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer C.A.J.A. PASCHE	Authorized officer

CLAIMS

1. Use of a retinoblastoma (Rb) protein for the control of the growth and/or replication of plant cells and plant viruses.

2. Use as claimed in claim 1 characterised in that the virus requires the integrity of an LXCXE amino acid motif in one of its proteins for the normal reproduction.

3. Use as claimed in claim 1 wherein the virus is a Geminivirus.

4. Use in accordance with claim 1 characterised in that the virus binds a retinoblastoma (Rb) protein in order to release a transcription factor.

5. A method of controlling the growth and/or replication of a plant cell or a plant virus within that cell, comprising the increase or decrease of the level and/or activity of a retinoblastoma protein in that plant cell.

6. A method as claimed in claim 5 characterised in that the level of protein is increased by direct application.

7. A method as claimed in claim 5 characterised in that the level of protein is increased by introduction of DNA or RNA encoding for its expression into the plant cell which it is desired to treat.

8. A method as claimed in claim 5, 6 or 7 wherein the protein is overexpressed.

9. A method of controlling the growth and/or replication of a plant cell or a plant virus comprising expressing an

Rb protein, or peptide fragment thereof that interacts with the LXCXE motif of the virus but does not affect the normal functioning of the cell, such as to inhibit cell growth or normal viral growth.

10. Recombinant nucleic acid encoding for expression of an Rb protein that has one or more characteristics of plant Rb protein not shared by animal Rb protein.

11. Nucleic acid as claimed in claim 10 characterised in that it comprises one or more characteristic regions that differ from known animal Rb protein nucleic acid.

12. Recombinant nucleic acid in the form of DNA or cRNA which encodes for a plant Rb protein having A and B pocket subdomains having a sequence with between 30% and 75% homology with human Rb protein.

13. Nucleic acid as claimed in claim 12 having a sequence with between 30% and 75% homology with p130 Rb retinoblastoma protein.

14. Nucleic acid as claimed in claim 12 or 13 characterised in that it has from 50% to 64% homology with animal or p130 Rb retinoblastoma protein.

15. Nucleic acid as claimed in any one of claims 12 to 14 encoding for the C706 amino acid of human Rb.

16. Nucleic acid as claimed in any one of claims 12 to 15 wherein the spacer sequence between the A and B pockets is not conserved with respect to animal Rb proteins.

17. Nucleic acid as claimed in claim 16 wherein the spacer sequence has less than 50% homology to the same

region found in animal retinoblastoma proteins.

18. Nucleic acid as claimed in any one of claims 12 to 17 having 80% or more homology with that of SEQ NO 2.

19. Nucleic acid as claimed in claim 18 wherein the homology is 90% or more.

20. Recombinant DNA comprising a sequence corresponding to SEQ ID No 1 bases 31 to 2079.

21. Recombinant DNA comprising a sequence corresponding to SEQ ID No 1 or corresponding RNA encoding for maize cDNA clone encoding ZmRb1 of SQ ID No 2.

22. Protein encoded by the recombinant DNA or RNA as claimed in any one of claims 12 to 21 or novel proteins derived from such DNA or RNA, and protein derived from naturally occurring DNA or RNA altered by mutagenic means.

23. Protein as claimed in claim 22 wherein the mutagenic means comprises mutagenesis using mutagenic PCR primers.

24. Anti-sense DNA or RNA of a gene encoding for a plant retinoblastoma protein, a gene which possesses the nucleic acid sequence as the one which is claimed in any one of the claims 10 to 21.

25. Vectors, cells, plants or animals comprising the DNA or RNA as claimed in any one of claims 12 to 22.

26. A method to control the growth and/or the proliferation of a vegetable cell or of a plant virus comprising the decrease of plant retinoblastoma protein

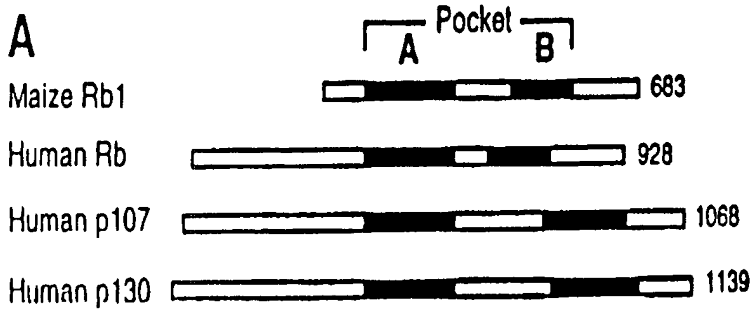
levels in the cell by incorporation to this cell of anti-sense DNA or RNA to the retinoblastoma protein.

27. cDNA encoding a protein as it is claimed in the claim 22.

28. A nucleic acid encoding a protein in which one or more of these sites are altered or deleted, making the protein more resistant to the phosphorilation and thus, to its functionality, for example, linking to E2F or similar.

29. An encoded protein by the nucleic acid which is described in claim 28.

A



B

Pocket A

```
Maize Rb1    89 TPVISAMTTA KWLREVISSL FDKPSSKLQQ FLSSCDRDLT NAVTERVSIV LEAIFPTKSS ANRGVSLGLN CANAFDIPWA EARKVEASKL 180
Xenopus Rb  348 TPVRGAMNTV QQLMVILSSA NDKPPDTLDS YFSNCTVNPK TKITDRIEHF GHVF.....K EKFASSVGQ. ..ACAEI..G YQRYKLGVCL 427
Chicken Rb  362 TPVRAAMNTI QQLMMILNSA TDKPSDTLIA YFNNCTVNPE DSILKRVECL GHIF.....K KKFAEAVGQ. ..GCAEI..G SQRYQLGVRL 441
Mouse Rb    367 TPVRTVMNTI QQLMVILNSA SDQPSENLIS YFNNCTVNPK ENILKRVKDV GHIF.....K EKFANAVGQ. ..GCVDI..G VQRYKLGVRL 446
Human Rb    373 TPVRTVMNTI QQLMMILNSA SDQPSENLIS YFNNCTVNPK ESILKRVKDI GYIF.....K EKFAKAVGQ. ..GCVEI..G SQRYKLGVRL 452
Human p107  385 TPVASATQSV SRLQSIVAGL KNAPSDQLIN IFESCVRNPV ENIMKILKGI GETF.....C QHYTQSTDEQ PGSHIDF..A VNRLKLAEIL 467
Human p130  417 TPVSTATHSL SRLHTMLTGL RNAPSEKLEQ ILRTCSRDPT QAIANRLKEM FEIY.....S QHF.QPDEDF SNCAKEI..A SKHFRFAEML 498

            180 YYRVLEAICR AELQNSNVNN LTPLLSNERF HRCLIACSAD LVLATHK... .........TV IMMFPAVLES TGLTAFDLSK IIENFVRHEE 259
            428 YYRVMEAILK TEEERLSVHN FSKLLNNDIF HICLLACAVE VVVASYARNA SQAY.CSSGT NLSFPWILRA FELKAFDFYK VIECFIKAEP 516
            442 YYRVMESMLK SEEERLSVHN FSKLLNDNIF HTSLLACALE IVMATYGRTA SQSDGTSAET DLSFPWILNV FDLKAFDFYK VIESFIKVEP 531
            447 YYRVMESMLK SEEERLSIQN FSKLLNDNIF HMSLLACALE VVMATYSRST LQHLD..SGT DLSFPWILNV LNLKAFDFYK VIESFIKVEA 534
            453 YYRVMESMLK SEEERLSIQN FSKLLNDNIF HMSLLACALE VVMATYSRST SQNLD..SGT DLSFPWILNV LNLKAFDFYK VIESFIKAEG 540
            468 YYKILETVMV QETRRLHGMD MSVLLEQDIF HRSLMACCLE IVLFAYSSPR .......... ..TFPWIIEV LNLQPFYFYK VIEVVIRSEE 545
            499 YYRVLESVIE QEQKRLGDMD LSGILEQDAF HRSLLACCLE VVTFSYKPPG .......... ..NFPFITEI FDVPLYHFYK VIEVFIRAED 576

            260 TLPRELKRHL NSLEEQLLES MAWEKGSSLY 289
            517 SLTSNMIKYL ERCEHQIMEC LAWQSDSPLF 546
            532 SLTRDMIKHL ERCEHRIMES LAWQSDSPLF 561
            535 NLTREMIKHL ERCEHRIMES LAWLSDSPLF 564
            541 NLTREMIKHL ERCEHRIMES LAWLSDSPLF 570
            546 GLSRDMVKHL NSIEEQILES LAWSHDSALW 575
            577 GLCREVVKHL NQIEEQILDH LAWKPESPLW 606
```

FIG. 1

2/3

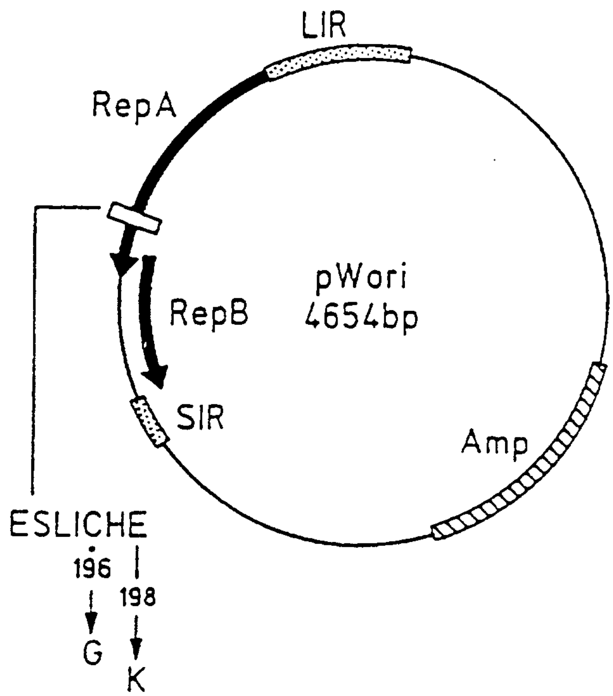
Pocket B

```
Maize Rb1    405 NEKCADVTIH IFFSKILKLA AIRIRNLCER VQC..VEQTE RVYNVFKQIL EQQTTLFFNR HIDQLILCCL YGVAKVCQLE LTFREILNNY  492
Xenopus Rb   614 QQKSTSLS.. LFYKKVYLLA YKRLSSLCSS LLSDHPELEQ VIWTLLQHTL QQEYELMRDR HLDQIMMCSM YGICKAKNID LRFKTIVTAY  701
Chicken Rb   630 QKPQKSTSLS LFYKKVFRLA YLRLHTLFFR LLSEHPDLEP LIWTLFQHTL QNESELMRDR HLDQIMMCSM YGICKVKNVD LRFKTIVSAY  719
Mouse Rb     632 QKPLKSTSLA LFYKKVYRLA YLRLNTLCAR LLSDHPELEH IIWTLFQHTL QNEYELMRDR HLDQIMMCSM YGICKVKNID LKFKIIVTAY  721
Human Rb     639 QKPLKSTSLS LFYKKVYRLA YLRLNTLCER LLSEHPELEH IIWTLFQHTL QNEYELMRDR HLDQIMMCSM YGICKVKNID LKFKIIVTAY  728
Human p107   780 NRPKRTGSLA LFYRKVYHLA SVRLRDLCLK LDVSN.ELRR KIWTCFEFTL VHCPDLMKDR HLDQLLLCAF YIMAKVTKEE RTFQEIMKSY  868
Human p130   828 NRPRKTSSLS LFFRKVYHLA AVRLRDLCAK LDISD.ELRK KIWTCFEFSI IQCPELMMDR HLDQLLMCAI YVMAKVTKED KSFQNIMRCY  916

                 493 KREAQCKPEV FSSIYI.... .......... .......... .......... .......... ...GSTNRNG VLVSRHVGII  525
                 702 KGLTNTNQET FKHVLIR... .......... .......... .......... .......... .......... ..DGQHDSII  726
                 720 KELPNTNQET FKRVLIR... .......... .......... .......... .......... .......... ..EEQYDSII  744
                 722 KDLPHAAQET FKRVLIR... .......... .......... .......... .......... .......... ..EEEFDSII  746
                 729 KDLPHAVQET FKRVLIK... .......... .......... .......... .......... .......... ..EEEYDSII  753
                 869 RNQPQANSHV YRSVLLK... .......SIP REVVAYNKNI NDDFEMID.. .......... .....CDLED ATKTPDCSSG PVKEERSDLI  931
                 917 RTQPQARSQV YRSVLIKGKR KRRNSGSSDS RSHQNSPTEL NKDRTSRDSS PVMRSSSTLP VPQPSSAPPT PTRLTGANSD MEEEERGDLI 1006

                 526 TFYNEVFVPA AKPFLV  541
                 727 VFYNLVFMQK LKSHIL  742
                 745 VFYNLVFMQK LKTNIL  760
                 747 VFYNSVFMQR LKTNIL  762
                 754 VFYNSVFMQR LKTNIL  769
                 932 KFYNTIYVGR VKSFAL  947
                1007 QFYNNIYIKQ IKTFAM 1002
```

FIG. 1


LIR
RepA
RepB
pWori
4654bp
SIR
Amp
ESLICHE
196
198
G
K

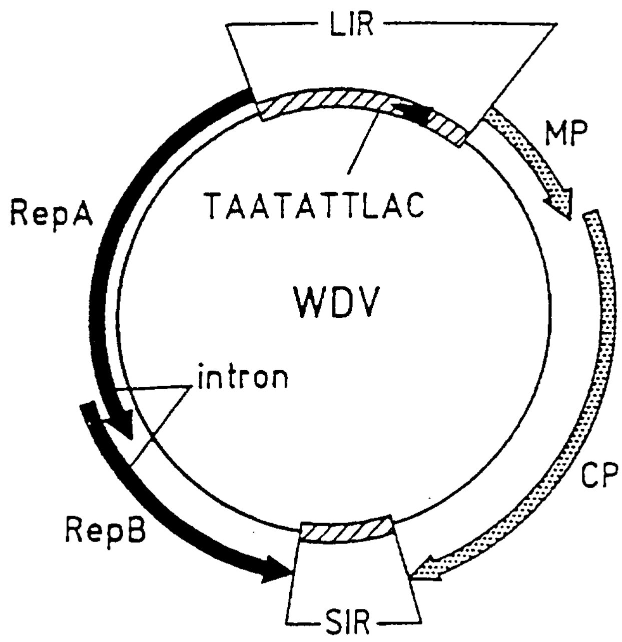

LIR
MP
RepA
TAATATTLAC
WDV
intron
CP
RepB
SIR

Fig. 2

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 97/03070

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/29 C12N15/82 C12N15/11 C12N5/10 C07K14/415

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SHEN B. ET AL.: "Partial sequencing and mapping of clones from two maize cDNA libraries" PLANT MOLECULAR BIOLOGY, vol. 26, no. 4, November 1994, pages 1085-1101, XP002042536 see the whole document & "AC T18395" EMBL DATABASE, 23 April 1994, HEIDELBERG, see the whole document ---	10-21, 25,27
X	GRAFI G. ET AL.: "AC U52099" EMBL DATABASE, 26 April 1996, HEIDELBERG, XP002042537	10-26
Y	see the whole document --- -/--	1-9,27

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

2

Date of the actual completion of the international search	Date of mailing of the international search report
6 October 1997	23.10.97

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Kania, T

International Application No
PCT/EP 97/03070

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	QIAN Y ET AL: "BIOLOGICAL FUNCTION OF THE RETINOBLASTOMA PROTEIN REQUIRES DISTINCT DOMAINS FOR HYPERPHOSPHORYLATION AND TRANSCRIPTION FACTOR BINDING" MOLECULAR AND CELLULAR BIOLOGY, vol. 12, no. 12, pages 5363-5372, XP000615356 see the whole document ---	28,29
X	WO 95 06661 A (RES DEV FOUNDATION ;FUNG YUEN KAI (US)) 9 March 1995 * see especially p.31, first par. * ---	28,29
Y	XIE Q. ET AL.: "Identification and analysis of a retinoblastoma binding motif in the replication protein of a plant DNA virus: requirement for efficient viral replication" THE EMBO JOURNAL, vol. 14, no. 16, 15 August 1995, pages 4073-4082, XP002042538 cited in the application * see the whole document, esp. pp. 4079/80 * ---	1-9,27
A	WO 95 07708 A (UNIV CALIFORNIA ;CANJI INC (US)) 23 March 1995 see the whole document ---	1-29
A	WO 92 05272 A (UNIV CALIFORNIA) 2 April 1992 see the whole document ---	24,26
A	COLLIN S. ET AL.: "The two nonstructural proteins from wheat dwarf virus involved in viral gene expression and replication are retinoblastoma-binding proteins" VIROLOGY, vol. 219, no. 1, 1 May 1996, pages 324-329, XP002042539 * see the whole document, esp. p.325, right col. * ---	1-29
A	SONI R. ET AL.: "A family of cyclin D homologs from plants differentially controlled by growth regulators and containing the conserved retinoblastoma protein interaction motif" THE PLANT CELL, vol. 7, no. 1, January 1995, pages 85-103, XP002042540 cited in the application * see the whole document, esp. p.97, right col. * --- -/--	1-29

2

ıational Application No
PCT/EP 97/03070

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	XIE Q. ET AL.: "Plant cells contain a novel member of the retinoblastoma family of growth regulatory proteins" THE EMBO JOURNAL, vol. 15, no. 18, 16 September 1996, pages 4900-4908, XP002042541 see the whole document ---	1-29
P,X	GRAFI G. ET AL.: "A maize cDNA encoding a member of the retinoblastoma protein family: involvement in endoreduplication" PNAS, U.S.A., vol. 93, no. 17, 20 August 1996, pages 8962-8967, XP002042542 see the whole document -----	1-29

2

International application No
PCT/EP 97/03070

INTERNATIONAL SEARCH REPORT

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically
Please see Further Information sheet enclosed.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA210

OBSCURITIES:
Claims 28 and 29 are formulated in a very inconcise manner. Consequently, the subject matter claimed was interpreted as follows and searched :

Claim 28: "A nucleic acid encoding a protein in which one or more sites are altered or deleted, making the protein more resistant to the phosphorilation and thus to it's functionality, for example, linking to E2F or similar "

Claim 29 : Unchanged.

Meaningful search not possible on the basis of all claims:
In claim 18 Seq ID 2 was read as Seq ID 1.

INTERNATIONAL SEARCH REPORT

Information on patent family members

ɿational Application No
PCT/EP 97/03070

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9506661 A	09-03-95	AU 7642694 A	22-03-95
		CA 2170605 A	09-03-95
		CN 1133595 A	16-10-96
		EP 0716660 A	19-06-96
		JP 9502183 T	04-03-97
		ZA 9406595 A	28-02-96
WO 9507708 A	23-03-95	AU 7834994 A	03-04-95
		ZA 9407065 A	03-05-95
WO 9205272 A	02-04-92	CA 2091759 A	18-03-92
		EP 0556231 A	25-08-93
		JP 6501160 T	10-02-94
		US 5324654 A	28-06-94